University of Minnesota Law School
Scholarship Repository

Minnesota Law Review | Minnesota Law Review Collections

2017

Stranger Than Science Fiction: The Rise of A.I. Interrogation in the Dawn of Autonomous Robots and the Need for an Additional Protocol to the U.N. Convention Against Torture

Amanda McAllister

Follow this and additional works at: https://scholarship.law.umn.edu/mlr

Part of the Law Commons

Recommended Citation

Amanda McAllister, *Stranger Than Science Fiction: The Rise of A.I. Interrogation in the Dawn of Autonomous Robots and the Need for an Additional Protocol to the U.N. Convention Against Torture*, 101 2527 (2017), *available at* https://scholarship.law.umn.edu/mlr/180.

This Article is brought to you for free and open access by the University of Minnesota Law School. It has been accepted for inclusion in Minnesota Law Review collection by an authorized administrator of the Scholarship Repository. For more information, please contact bkeele@umn.edu.

Note

Stranger than Science Fiction: The Rise of A.I. Interrogation in the Dawn of Autonomous Robots and the Need for an Additional Protocol to the U.N. Convention Against Torture

*Amanda McAllister**

Amidst a bloody civil war, a group of rebel insurgents have stolen critical information regarding the infrastructure of the government's head military base; their intent is to identify structural weaknesses and coordinate a massive attack that will result in hundreds of thousands of casualties. The commander of the military manages to identify a member of the insurgent group while she is posing as a diplomat on a consular assistance mission. He takes her into custody and accuses her of treason. Desperate to prevent the impending attack, he interrogates her while she is being tortured in an attempt to obtain critical information on the insurgent group's activities.[1]

But there is one catch: the diplomat is Princess Leia, the commander, Darth Vader. And the torturer? An autonomous droid.[2]

* J.D. Candidate 2017, University of Minnesota Law School. I am especially thankful to Professor Fionnuala Ní Aoláin for her thought-provoking questions and invaluable guidance while drafting this Note. I also want to extend my gratitude to Jason Crider, Griffin Ferry, Mandi Janikowski, Megan Manion, and Hannah Nelson for their insights and feedback. Finally, a heartfelt thank you to my husband, Tyler, for his endless encouragement and love. Copyright © 2017 by Amanda McAllister.

1. Such an act is a violation of international human rights, international humanitarian law, and customary international law. *The Legal Prohibition Against Torture*, HUM. RTS. WATCH (Mar. 11, 2003), https://www.hrw.org/news/2003/03/11/legal-prohibition-against-torture#laws (explaining that torture is a violation of the Universal Declaration of Human Rights, the United Nations Convention Against Torture, international humanitarian law, and has become a jus cogens (non-derogable) norm and, as such, a violation of customary international law).

2. STAR WARS: EPISODE IV – A NEW HOPE (Lucasfilm 1977). The interro-

Who then is responsible for the actions of the robot, and consequently, the human rights violation? Is the system merely applying preprogrammed actions or making autonomous decisions? How does that implicate the responsibility of its designers or operators?[3] If the robot is autonomous, would it bear responsibility? Can an individual be held liable for the autonomous decisions of a non-human some*thing* else?

While such a nightmarish scenario may seem to belong in a galaxy far, far away, the interrogative dimensions of the torture *Star Wars* depicted might not be so distant after all. Many predict the rise of autonomous weapons systems to be on the horizon, with robotics technology positioned as the next Revolution in Military Affairs (RMA).[4] Indeed, drones and other limited autonomous robots are already operating on the battlefield, and countries with their eyes set on autonomous systems are spending billions on their development.[5] Artificial intelligence

gator droid is designed for torture and deliberately frightening in appearance, armed with syringes, flesh peelers, and bone fragmenters. SIMON BEECROFT, STAR WARS CHARACTER ENCYCLOPEDIA 94 (2011). It is "completely without pity." *Id.* For a broader discussion about threat of torture, see Gäfgen v. Germany, 2010 Eur. Ct. H.R. 1, 16, http://hudoc.echr.coe.int/eng?i=001-99015 ("Legal qualification of threats of torture."); GARY D. SOLIS, THE LAW OF ARMED CONFLICT: INTERNATIONAL HUMANITARIAN LAW IN WAR 437–85 (2010) (discussing how the laws around threat of torture are unsettled).

3. *The Ethics of Autonomous Weapons Systems*, U. PA. L. SCH., https://www.law.upenn.edu/institutes/cerl/conferences/ethicsofweapons (last visited May 12, 2017).

4. Ronald Arkin, *Lethal Autonomous Systems and the Plight of the Noncombatant*, 137 AISB Q. 4, 4 (2013); *see also* Chantal Grut, *The Challenge of Autonomous Lethal Robotics to International Humanitarian Law*, 18 J. CONFLICT & SECURITY L. 5, 6–7 (2013) ("[W]hile autonomous killer robots might sound like science fiction, they are much closer to reality than is often appreciated. See for example the drones capable of navigating their own flight paths, developed or being developed by Northrop Gunman [sic], Carnegie Mellon and MIT amongst others; the commissioning of research by the US Air Force into software which would allow drones to land and take-off autonomously; and 'swarming' drones, which by definition operate and move together autonomously."); Geoff Dyer, *U.S. Military: Robot Wars*, FIN. TIMES (Feb. 7, 2016), https://www.ft.com/content/849666f6-cbf2-11e5-a8ef-ea66e967dd44 ("For the Pentagon leadership, innovations like [autonomous drones] are centerpieces of a new wave of military technology that officials hope will keep the US ahead of China and Russia, whose heavy investments in recent years has [sic] closed the gap.").

5. *Losing Humanity: The Case Against Killer Robots*, HUM. RTS. WATCH (Nov. 19, 2012), https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots ("[T]he US Department of Defense has spent approximately $6 billion annually on the research and development, procurement, op-

has become increasingly adept at reading human emotional responses, making robot-enhanced interrogation capable of measuring physiological reactions related to truth-telling a real possibility.[6] Yet, "while robot interrogators offer time, cost, and effectiveness benefits, they also risk undermining fundamental rights."[7]

Internationally, there have been growing calls for bans or an international convention on autonomous robotic weapons.[8] The past three years, state parties have convened and debated such a ban at the United Nations (UN) Convention on Certain Conventional Weapons.[9] Yet, no legislation, treaty, or norms currently exist to address the unprecedented threat posed by autonomous robots to torture and cruel, inhuman, and degrading treatment prohibitions.

erations, and maintenance of unmanned systems for war, and that figure is likely to increase rapidly."); Mary Wareham, *Killer Robots: Why the World Should Ban Autonomous Weapons Systems*, HUM. RTS. WATCH (Nov. 28, 2014), https://www.hrw.org/news/2014/11/28/killer-robots-why-world-should-ban-autonomous-weapons-systems ("[S]everal precursors [of autonomous weapons] . . . are in development in the United States, China, Israel, Russia, South Korea, the United Kingdom, and other nations with high-tech militaries demonstrat[ing] the trend toward ever-increasing autonomy on land, in the air, and on or under the water.").

6. Kristen Thomasen, *Examining the Constitutionality of Robot-Enhanced Interrogation*, *in* ROBOT LAW 306, 315–16 (Ryan Calo et al. eds., 2016).

7. *Id.* at 329.

8. Matthew Bolton, *Debating Proposals for a Possible International Convention on Robotic Weapons*, INT'L COMM. FOR ROBOT ARMS CONTROL (Oct. 3, 2012), http://icrac.net/2012/10/debating-proposals-on-a-possible-international-convention-on-robotic-weapons.

9. In early 2015, the UN held a Convention on Certain Conventional Weapons (CCW) to debate a ban on lethal autonomous weapons systems. Currently, the United States delegation to the CCW has been noncommittal on a ban and United States policy permits the Department of Defense to pursue such a development. Steven Groves, *U.N. Conference Debating a Ban on Autonomous Weapons: Understanding Key Issues*, HERITAGE (Apr. 16, 2015), http://www.heritage.org/defense/report/un-conference-debating-ban-autonomous-weapons-understanding-key-issues. A U.S. Department of Defense Directive in 2012 required that, for now, human beings remain "in the loop" when decisions involving lethal force are made. *See infra* note 38. The directive contains significant loopholes, such as a waiver by high-level department officials and expiration without renewal in 2022. Bonnie Docherty, *U.S. Takes First Step on Fully Autonomous Weapons, but Stricter Controls Needed*, HUM. RTS. @HARVARD L. BLOG (Apr. 16, 2013), http://hrp.law.harvard.edu/staff/u-s-takes-first-step-on-fully-autonomous-weapons-but-more-needed.

Many have accused the laws of armed conflict of being backwards-looking and "one war behind" in its regulation of battle spaces.[10] Similarly, international human rights law runs the risk of being "one technological revolution" behind if it fails to foresee and prevent the widening gap for potential human rights violations likely to exist in a more fully realized autonomous future. Further, the dawn of autonomous warfare is not just any technological revolution; it is a paradigm shift and *qualitatively different* than anything of the past.[11] Therefore, it is "imperative that courts, state agencies, legislators, and designers consider and address the potential implications of this technology before it is widely deployed."[12]

While some have called for a total ban on the development of autonomous systems,[13] many consider that to be an unsustainable or impractical solution.[14] As such, neglecting to critically analyze crevices in coverage in the language of international law will allow for rhetorical circumvention based on hazy understandings of robots' growing autonomy. This in turn will risk impunity for grave human rights violations. Our own current visions of technology and the language we have ascribed to them are not static. An inevitable reformulation of linguistic approaches post-technological revolution risks rendering the language of the foremost treaty on torture and cruel, inhuman and degrading treatment prohibitions, the Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or

10. *See* FRANCOISE BOUCHET-SAULNIER, THE PRACTICAL GUIDE TO HUMANITARIAN LAW, at xvi (2014) ("[Inability] wrongly implies that States are unable to anticipate crimes and abuses they will later accept to commit, whether in the name of highly questionable or highly legitimate causes.").

11. *Autonomous Weapons – Q &A*, INT'L COMMISSION FOR THE RED CROSS (Nov. 10, 2014), https://www.icrc.org/en/document/autonomous-weapon-systems-challenge-human-control-over-use-force.

12. Thomasen, *supra* note 6, at 329.

13. *See, e.g.*, HUMAN RIGHTS WATCH, SHAKING THE FOUNDATIONS: THE HUMAN RIGHTS IMPLICATIONS OF KILLER ROBOTS 5 (2014) (outlining several nations calling for a ban); Bonnie Docherty, *The Trouble with Killer Robots*, FOREIGN POL'Y (Nov. 19, 2012), http://foreignpolicy.com/2012/11/19/the-trouble-with-killer-robots (arguing "governments should adopt an international prohibition on the development, production, and use" of such weapons).

14. *See, e.g.*, GARY D. SOLIS, THE LAW OF ARMED CONFLICT: INTERNATIONAL HUMANITARIAN LAW IN WAR 535 (2d ed. 2016); Kenneth Anderson & Matthew C. Waxman, *Laws and Ethics for Autonomous Weapons Systems: Why a Ban Won't Work and How the Laws of War Can* 2 (Am. Univ. Wash. Coll. of Law, Research Paper No. 2013–11).

Punishment (CAT), obsolete in the realm of autonomous interrogation.

This Note addresses the unprecedented effect autonomous robots will have on the scope of current prohibitions against torture and cruel, inhuman and degrading treatment and explains the need for a linguistic reexamination to close gaps in language likely to result from a shared post-human world.[15] This emergent space recognizes the increased prominence of post-humanism, or the ideology that moves away from the human as possessing an essential element and speaking to the co-evolutionary spiral of technology and humanity as intertwined, thus calling into question conceptions of autonomy and agency.[16]

Part I of this Note examines current international prohibitions on and rhetorical circumventions of torture as well as theoretical bases for scope limitations of torture prohibitions. Part II discusses how the lack of consensus regarding post-human visions of technology may create a vacuum of liability upon the introduction of autonomous robots into interrogative spaces and threaten the ability to hold those responsible for torture accountable. Part III proposes guidelines for an Additional Protocol to the CAT to specifically and unambiguously ban torture and other cruel, inhuman, or degrading treatment or punishment through mediums of automata or autonomous devices. Accordingly, this Note concludes that a specifically tailored prohibition is necessary to avoid being "one technological revolution" behind in the protection of one of the most fundamental human rights: the right to be free from torture.[17]

15. This Note does not engage in the important debate surrounding the moral and ethical questions inherent in creating and utilizing such systems, nor does it examine the arguments that autonomous weapons systems are a technological impossibility. Further, while there is an important debate regarding the validity and workability of complete bans on the development of autonomous weapons systems, this Note instead assumes the inevitability of the existence of such systems. This Note also will not discuss the significant constitutional issues to be assessed with regard to the right to silence and the right to privacy in the context of interrogation performed by autonomous robots. *See, e.g.*, Thomasen, *supra* note 6.

16. N. KATHERINE HAYLES, HOW WE BECAME POSTHUMAN: VIRTUAL BODIES IN CYBERNETICS, LITERATURE, AND INFORMATICS 2–4 (1999).

17. G.A. Res. 217 (III) A, art. 5, Universal Declaration of Human Rights (Dec. 10, 1948).

I. AUTONOMOUS ROBO-INTERROGATORS AND THE THREAT POSED TO THE SCOPE OF TORTURE PROHIBITIONS

Prohibitions on torture have proliferated extensively since the Second World War.[18] However, the integration of autonomous robots into torture and interrogation spaces threatens to render previous protections stagnant and leave newly created gaps open to exploitation and, consequently, impunity. This Part will discuss the new and emerging technologies in domestic, battlefield, and interrogation spaces as well as the limitations constricting coverage, including rhetorical, cultural, and doctrinal components facing disjunction upon the introduction of autonomous robots.

Section A discusses the concept of autonomous weapons systems, limited autonomy in current use, and the projected development of autonomous robots. Further, it explores the technologies most likely to be utilized in a robo-interrogative scenario. Section B addresses constrictions to the scope of the prohibition, including the current, decaying language of the CAT and the infamous rhetorical evasion of its coverage in the U.S. "torture memos" controversy. Last, it considers theoretical limitations, including the cultural relativism of anthropomorphization[19] of robots and predictive conceptions of post-humanism.[20]

18. *See supra* note 1.

19. Anthropomorphism is the attribution of human traits or behaviors to a non-human:

> One approach to enhance people's acceptance of robots is the attempt to increase a robot's familiarity by using anthropomorphic (human-like) design and "human social" characteristics. This implies human-like parts of a robot's physical shape, the usage of facial expressions and other social cues, as well as natural humanlike interaction and communication . . . . An underlying assumption is that humans prefer to interact with machines in the same way they interact with other people.

Julia Fink, *Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction*, 7621 SOC. ROBOTICS 199, 199 (2012).

20. Post humanism provides a recalibrated lens with which to view the intertwining of humans and technology:

> As long as the human subject is envisioned as an autonomous self with unambiguous boundaries, the human-computer interface can only be parsed as a division between the solidity of real life on one side and the illusion of virtual reality on the other, thus obscuring the far-reaching changes initiated by the development of virtual technologies . . . . [W]hen the human is seen as part of a distributed system, the

A. ROBOTS: AUTONOMY, PROLIFERATION, AND INTEGRATION INTO INTERROGATION

This Section will discuss the current state of robotics development. It will also explain the projected development and current efforts underway to achieve true artificial intelligence. Last, it will explore theories on how robots would be utilized in interrogative spaces.

1. Autonomous Robots

It is critical to define what constitutes an "autonomous" robot as opposed to "automatic" one. First, robots typically require "three key components: 'sensors' that monitor the environment and detect changes in it, 'processors' or 'artificial intelligence' that decides how to respond, and 'effectors' that act upon the environment in a manner that reflects the decisions of the processors."[21] Then, autonomy depends on the robot's ability to "operate without any form of external control for an extended period."[22] Autonomous weapons would be "weapons with the capacity to utilize artificial intelligence to replicate human cognitive reasoning."[23] So, while many military weapons, such as drones, are still controlled by an operator, an autonomous[24] sys-

> full expression of human capability can be seen precisely to *depend* on the splice rather than being imperiled by it.

HAYLES, *supra* note 16, at 290; *cf.* Francesca Ferrando, *Posthumanism, Transhumanism, Antihumanism, Metahumanism, and New Materialisms: Differences and Relations*, 8 EXISTENZ 26, 26 (2013) (arguing posthuman has become a confusing umbrella term).

21. Michael Kolb, Soldier and Robot Interaction in Combat Environments 8 (2012) (unpublished Ph.D. dissertation, University of Oklahoma) (on file with author).

22. *Id.* at 9; *cf.* Merel Noorman & Deborah G. Johnson, *Negotiating Autonomy and Responsibility in Military Robots*, 16 ETHICS INFO. TECH. 52 (2014) (arguing the concept of autonomy is used metaphorically).

23. Geoffrey S. Corn, Autonomous Weapons Systems: Managing the Inevitability of "Taking the Man out of the Loop" (June 14, 2014), http://ssrn.com/abstract=2450640 (citation omitted).

24. Yet, there are many stops on the spectrum of autonomy for a weapons system. For example, remote-controlled systems have virtually no autonomy, as they are incapable of acting without an operator. *See* Kolb, *supra* note 21, at 24. Tele-operated systems may have one or more integrated sensors used for navigation and other tasks. *Id.* Semi-autonomous systems would use "onboard sensing and processing" to control some of their functions. *Id.* Many unmanned aerial vehicles, such as the Grey Eagle UAV, are semi-autonomous systems. *Id.*

tem would be able to control and make decisions for itself.[25] Truly autonomous systems would be self-governing and incorporate strong artificial intelligence capable of "choosing their own course of action to reach a desired goal."[26]

2. Development of Autonomous Systems

This Section explores the current state of artificial intelligence and autonomy in use on the battlefield as well as the projected development of autonomous weapons on the battlefield, domestic, and public spaces.

a. Limited Autonomous Systems and Projected Development

Robotic systems are "widely present in the modern battlefield providing intelligence gathering, surveillance, reconnaissance, target acquisition, designation and engagement capabilities."[27] The Samsung SGR-A1 surveillance and security guard robot on the border of North and South Korea is capable of selecting and engaging targets and delivering lethal force without intervention by a human operator.[28] There are also the human-supervised U.S. autonomous systems—"the Aegis at sea and the Patriot on land"—capable of defending "against short-notice missile attacks."[29] Other examples of such systems with limited autonomy are the Phalanx system, "capable of autonomously performing its own search, detect, evaluation, track, engage and kill assessment functions," and "Israel's Iron Dome [anti-missile] system."[30] In addition, there is the tele-operated iRobot Packbot series, the unmanned ground vehicle (UGV) Foster-Miller TALON series, and the AeroVironment RQ-11 Raven, an unmanned aerial vehicle (UAV) capable of complete autonomous flight via GPS waypoint navigation.[31]

In the realm of artificial intelligence, one of the most famous example of a computer displaying high levels of intelligence is the IBM computer Deep Blue.[32] On May 11, 1997, Deep

25. Grut, *supra* note 4, at 5.
26. SOLIS, *supra* note 14, at 537–38; Grut, *supra* note 4, at 6.
27. Arkin, *supra* note 4.
28. SOLIS, *supra* note 14.
29. *Id.* at 535–36.
30. *Id.* at 536; Arkin, *supra* note 4.
31. Kolb, *supra* note 21, at 10–12.
32. *Icons of Progress: Deep Blue*, IBM, http://www-03.ibm.com/ibm/history/

Blue beat the world chess champion in a match that lasted several days.[33] It demonstrated the ever-increasing ability of computers to handle complex calculations, perform risk analysis, and identify trends. Deep Blue also stands as an example for the potential of computers to possess intentionality.[34]

Truly autonomous systems have not been developed yet.[35] Yet, the advantages of such systems, including "enhanc[ing] situational awareness, reduc[ing] human workload, improv[ing] mission performance and minimiz[ing] overall risk to civilian and military personnel, and all at a reduced cost" will push nations to continue their development.[36] Regardless of how imminent true autonomy is, governments and the United Nations are preparing for the possibility and developing more limited-autonomy systems.[37] The U.S. Department of Defense issued the world's first policy pronouncement on autonomous weapons

ibm100/us/en/icons/deepblue (last visited May 12, 2017).

33. *Id.*

34. *The Chinese Room Argument: Intentionality*, STAN. ENCYCLOPEDIA PHIL. (Apr. 9, 2014), https://plato.stanford.edu/entries/chinese-room/#5.2 ("Intentionality is the property of being about something, having content . . . . Beliefs and desires are intentional states: they have propositional content . . . . [A]pplie[d] to computers[,] . . . [w]e can interpret the states of a computer as having content, but the states themselves do not have original intentionality."). After Deep Blue's success, IBM built Watson, a computer that beat two of the most successful human players of *Jeopardy!* while also demonstrating a substantial advancement in programming from Deep Blue. *See* Joe Best, *IBM Watson: The Inside Story of How the Jeopardy-Winning Supercomputer Was Born, and What It Wants To Do Next*, TECHREPUBLIC (Sept. 10, 2013), http://www.techrepublic.com/article/ibm-watson-the-inside-story-of-how-the-jeopardy-winning-supercomputer-was-born-and-what-it-wants-to-do-next. Watson was capable of processing and reasoning from natural language and performing accurate processing of its massive supply of information. Daniel C. Dennett, *When HAL Kills, Who's To Blame? Computer Ethics*, *in* HAL'S LEGACY: 2001'S COMPUTER AS DREAM AND REALITY 352 (David G. Stork ed., 1997). While agreeing "skeletal versions of human beliefs and desires" are not sufficient to meet mens rea requirement of legal culpability, he states, "it is hard to see what is missing." *Id.*

35. Grut, *supra* note 4, at 6.

36. SOLIS, *supra* note 14, at 536.

37. There has been significant public backlash to the development of autonomous systems, as well. More than 1000 robotics researchers and experts signed an open letter calling for a ban on "offensive autonomous weapons," including Steve Wozniak, Demis Hassabis, Stephen Hawking, and Elon Musk. Samuel Gibbs, *Musk, Wozniak and Hawking Urge Ban on Warfare AI and Autonomous Weapons*, THE GUARDIAN (July 27, 2015), http://www.theguardian.com/technology/2015/jul/27/musk-wozniak-hawking-ban-ai-autonomous-weapons.

systems in 2012, stating that, for the term of the directive, humans must remain "in the loop," or, the last point of authorization in any decision involving lethal force, in the application of lethal force by autonomous weapons systems.[38] The directive must be reissued, cancelled, or certified by November 21, 2017, or it will expire on November 21, 2022.[39]

In 2015 and 2016, the UN held their annual Convention on Certain Conventional Weapons to debate a ban on lethal autonomous weapons systems, fearing a "robotic arms race" leading to deployment of lethal weapons incapable of compatibility with the laws of armed conflict and international law.[40]

There are no multinational treaties, UN conventions, or other policy pronouncements that regulate this emerging space. Yet, as this "new age of machines empowered to make decisions about life and death . . . dawn[s] a[nd] research in artificial intelligence advances," discussion and debate on autonomous weapons have drastically increased.[41]

b. Robots and Autonomous Systems in Domestic and Public Spaces

Proliferation of robots and other autonomous systems is unlikely to be restricted to the battlefield.[42] In the United States, the National Robotics Initiative, supported by multiple agencies of the federal government, released a plan to fund research on robots, including the field of Artificial Intelligence

38. U.S. Dep't of Def., Directive 3000.09, Autonomy in Weapons Systems (2012), http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf.

39. *Id.*

40. *See supra* note 9 and accompanying text.

41. Andrew Hanna, *Killer Robots Await Trump's Verdict*, POLITICO (Dec. 25, 2016), http://www.politico.com/story/2016/12/drones-military-technology-trump-232933.

42. The Ministry of Economy, Trade and Industry of Japan (METI) designated robots as one of "the seven most critical industrial fields" in its Strategy for Creating New Industries. Naho Kitano, Animism, Rinri, Modernization; The Base of Japanese Robotics 4 (Jan. 5, 2007) (unpublished manuscript) (on file with author). METI released *New Robot Strategy: Vision, Strategy, Action Plan* in 2015, which claims Japan's status as a "Robotics Superpower" and warns that "Japan will be isolated from the rest of the world in the field of robots" if Japan lags behind in robot development. MINISTRY OF ECON., TRADE & INDUS., NEW ROBOT STRATEGY: JAPAN'S ROBOT STRATEGY: VISION, STRATEGY, ACTION PLAN 2, 5 (2015). METI further urged a "major innovation in robotics including robot technology and robot utilization system[s]" to serve as effective tools in Japan's social and economic challenges. *Id.* at 6.

(AI),[43] with the goal of accelerating development and use of robots in the United States that "work beside, or cooperatively with, people."[44] In October 2016, the White House released the report *Preparing for the Future of Artificial Intelligence* to address opportunities and challenges of such systems and to provide guidelines on the development and incorporation of artificial intelligence in various sectors to contribute to economic growth.[45] On November 30, 2016, Senator Ted Cruz chaired the Senate Commerce Subcommittee on Space, Science and Competitiveness hearing, "The Dawn of Artificial Intelligence," the first-ever congressional hearing dedicated to AI. There, he declared we are "on the verge of a new technological revolution thanks to rapid advances in processing power."[46]

In addition, demand for robots worldwide has increased dramatically.[47] Market researcher Forrester released a report predicting a greater than 300% increase in artificial intelligence investments.[48] Indeed, this past year alone, we have seen widespread domestic and commercial use of robots, including unmanned aerial drones.[49]

Self-driving cars have moved from a dream of the future to a reality: the 2017 Kia Forte touts its "autonomous emergency

43. NAT'L ROBOTICS INITIATIVE, THE REALIZATION OF CO-ROBOTS ACTING IN DIRECT SUPPORT OF INDIVIDUALS AND GROUPS 2 (Oct. 4, 2012), http://www.nsf.gov/pubs/2012/nsf12607/nsf12607.pdf.

44. *Id.* Anticipated funding is $40 million to $50 million per year. *Id.*

45. NAT'L SCI. & TECH. COUNCIL, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE (2016) (addressing regulation of AI); *see also* Ed Felten & Terah Lyons, *The Administration's Report on the Future of Artificial Intelligence*, WHITE HOUSE BLOG (Oct. 12, 2016), https://www.obamawhitehouse.archives.gov/blog/2016/10/12/administrations-report-future-artificial-intelligence.

46. Press Release, Ted Cruz, Sen. Cruz Chairs First Congressional Hearing on Artificial Intelligence (Nov. 30, 2016) (on file with author). Senator Cruz argued, "Ceding leadership in developing artificial intelligence to China, Russia and other foreign governments will not only place the United States at a technological disadvantage but it could also have implications for national security." *Id.* He concluded by declaring that while "[t]his is the first congressional hearing on artificial intelligence . . . it will not be the last." *Id.*

47. *Id.*

48. Gil Press, *Forrester Predicts Investment in Artificial Intelligence Will Grow 300% in 2017*, FORBES (Nov. 1, 2016), http://www.forbes.com/sites/gilpress/2016/11/01/forrester-predicts-investment-in-artificial-intelligence-will-grow-300-in-2017.

49. Drew Prindle, *The Best Drone You Can Buy*, DIGITAL TRENDS (Oct. 19, 2016), http://www.digitaltrends.com/cool-tech/best-drones.

braking systems" as a solution to distracted driving.[50] As of October 19, 2016, all Tesla vehicles have the requisite hardware for full self-driving capabilities.[51] Ford has announced it will release self-driving cars by 2021 to enhance mobility, announcing itself as a "company of the future, not of the past."[52] Golf carts and scooters are also implementing autonomous car software and sensor configuration.[53] Several states have passed autonomous vehicle legislation to permit testing and use of self-driving cars and trucks on roads.[54]

Robots are increasingly being integrated into the workforce and replacing human workers.[55] For example, Fukoku Mutual Life Insurance is utilizing IBM's Watson Explorer, capable of analysis and interpretation of text, image, audio and visual da-

50. Kia Motors Am., *The New 2017 Kia Forte – Car Karaoke with Autonomous Emergency Breaking*, YOUTUBE (Sept. 7, 2016), https://www.youtube.com/watch?v=Z3aBucvZ_Vg.

51. *All Tesla Cars Being Produced Now Have Full Self-Driving Hardware*, TESLA: BLOG (Oct. 19, 2016), https://www.tesla.com/blog/all-tesla-cars-being-produced-now-have-full-self-driving-hardware (describing the hardware specifications). Tesla claims this will be "substantially safer" than a human driver. *Id.*

52. Daniel Newman, *Autonomous Cars: The Future of Mobility*, FORBES (Sept. 27, 2016), http://www.forbes.com/sites/danielnewman/2016/09/27/autonomous-cars-the-future-of-mobility. BMW is likely on the same timeline as Ford. *Id.*

53. *Driverless Cars, Golf Carts, Now Joined by Autonomous Scooter*, SCI. DAILY (Nov. 8, 2016), https://www.sciencedaily.com/releases/2016/11/161108120609.htm.

54. Press Release, Mich. Governor Rick Snyder, Gov. Rick Snyder Signs Landmark Legislation To Allow Operation of Autonomous Vehicles on Michigan Roadways (Dec. 9, 2016), http://www.michigan.gov/snyder/0,4668,7-277--399173--,00.html. Some robotics experts predict that that car ownership and use as we know it will be fundamentally changed by the autonomous future. Alex Nishimoto, *Robotics Expert Predicts Kids Born Today Will Never Drive a Car*, MOTOR TREND (Jan. 3, 2017), http://www.motortrend.com/news/robotics-expert-predicts-kids-born-today-will-never-drive-car.

55. Some fear the impact on industries could lead to displacement of jobs spread unevenly across the economy, leading to an exacerbation of wealth inequality. Arjun Kharpal, *AI Could Boost Productivity but Increase Wealth Inequality, the White House Says*, CNBC (Dec. 21, 2016), http://www.cnbc.com/2016/12/21/ai-could-boost-productivity-but-increase-wealth-inequality-the-white-house-says.html; *see also* Mark B. Solomon, *Autonomous Heavy-Duty Trucks Threaten Jobs of Nearly 1.7 Million Drivers, White House Says*, D.C. VELOCITY (Jan. 3, 2017), http://www.dcvelocity.com/articles/20170103-autonomous-heavy-duty-trucks-threaten-jobs-of-nearly-17-million-drivers-white-house-says.

ta, to calculate payouts to policyholders.[56] A Nomura Research Institute report has claimed that robots could perform half of the jobs in Japan by 2035.[57] The technology is rapidly evolving, and necessitates responses from law and policy.

3. Robots in Policing Spaces

Even with the increasing amount of robots in domestic spaces, the thought of robots being used by police or security forces in interrogation spaces sounds like something out a science fiction film. Yet, robots have already been incorporated into domestic policing in the United States. During the 2016 shooting of Dallas police officers, law enforcement used a teleoperated robot armed with C-4 plastic explosives to remotely detonate near and kill the sniper who ambushed five officers.[58] This surprised many, as "the public is not widely aware that police departments have robots."[59]

Police departments have also used unmanned aircrafts.[60] Several states have responded with requiring police to obtain warrants prior to any aerial surveillance or by limiting the ability to retain collected data.[61] No such responses have occurred regarding ground-based robots, although police are using robots to defuse explosives and deliver cell phones to facilitate negotiations in the expansion of this uncharted territory.[62]

Further, the use of robots by domestic agencies is only likely to grow. The FBI, sheriff's departments, and federal and local agencies all are in possession of robots, and often from a federal program that has transferred $6 billion in excess U.S. military equipment to local law enforcement agencies since the

56. Justin McCurry, *Japanese Company Replaces Office Workers with Artificial Intelligence*, THE GUARDIAN (Jan. 5, 2017), https://www.theguardian.com/technology/2017/jan/05/japanese-company-replaces-office-workers-artificial-intelligence-ai-fukoku-mutual-life-insurance.

57. *Id.*

58. Kevin Sullivan et al., *Dallas Police Used a Robot To Kill. What Does That Mean for the Future of Police Robots?*, WASH. POST (July 21, 2016), https://www.washingtonpost.com/national/dallas-police-used-a-robot-to-kill-what-does-that-mean-for-the-future-of-police-robots/2016/07/20/32ee114e-4a84-11e6-bdb9-701687974517_story.html.

59. *Id.*

60. *Id.*

61. *Id.*

62. *Id.*

1990s.[63] As such, the proliferation of autonomous weapons systems in armed conflict and military spaces will continue being reflected in domestic policing.

Since the Dallas shootings, many have expressed concern over weaponizing robots for police purposes.[64] Indeed, the "absence of law and policy . . . has the potential to lead to overuse of machines that can be used to injure or kill suspects."[65] These concerns are compounded in the context of interrogation.

4. Robots in Interrogative Spaces

Now, it may be difficult to imagine what legitimate purposes robots could serve in lawful interrogative spaces. However, with increasing developments in human-computer interaction research and physiological measurement devices, combined with the inability of humans to act as effective lie detectors and the traditional reliance on technology to enhance human capacities, such a space is "not out of the realm of possibility."[66] Indeed, autonomous interviewers would permit governments and agencies reliable, low-intrusive, and cost-effective means of interviewing large groups of people quickly or detecting weapons at common security points.[67]

This Section will discuss what a robo-interrogator would be, the technology it would likely be equipped with, and existing precursors.

a. A Robo-Interrogator

A robot interrogator would be "any automated technology that examines an individual through questioning . . . for the

63. *Id.* "Since 2003, 280 law enforcement agencies have obtained at least 987 reconnaissance and bomb-disposal robots from the military . . . ." *Id.*

64. Brigitt Keller of the National Police Accountability Project argues that use of a robot for killing was equivalent to being summarily executed outside the Constitution. *See id.*

65. *Id.*

66. Kristen Thomasen, Liar, Liar, Pants on Fire! Examining the Constitutionality of Enhanced Robo-Interrogation 2 (2012) (unpublished manuscript), http://robots.law.miami.edu/wp-content/uploads/2012/01/Thomasen_CONSTITUTIONALITY-OF-ROBOT-INTERROGATION.pdf. "Humans are generally poor lie detectors (not usually accurate above 60%) . . . ." *Id.* at 1 n.2.

67. Simon Parkin, *The Police Are Recruiting Interrogation Bots*, HOW WE GET TO NEXT (May 20, 2016), https://howwegettonext.com/the-police-are-recruiting-interrogation-bots-e6fd68286ef3.

purpose of eliciting incriminating statements or confessions."[68] One scholar argues that "psychological interrogation techniques and behavioral lie-detection methods" are the "twin pillars" of interrogation.[69] As such, the robot interrogator would have to exceed the human interrogator in these capabilities to be considered desirable and incorporated into the interrogation space.[70]

In 1983, the CIA experimented with the first known robo-interrogator.[71] In the since declassified document "Interrogation of an Alleged CIA Agent," the CIA attempted the use of a microcomputer in interrogation through Analiza, a primitive artificial intelligence program which utilized sophisticated algorithms to formulate replies.[72] Significantly, "[b]ecause it retains in its memory the topics covered, it will have more surprises in store for [the subject] as the sessions continue; with time it becomes increasingly more knowledgeable about [its subject]."[73]

The program probed for its subject's vulnerabilities and exerted defense measures against hostility.[74] The report concludes by stating that "[a]s for [the subject], he is fortunate that should the probing get too discomforting, he will have an option that will not be available to him in a true overseas interview situation—he can stop the questions with a flick of the 'off' switch."[75] True robo-interrogators would provide no such comfort to the detainees.

b. Deception-Detection

A robot interrogator, with the combination of human-computer technology and sensors capable of monitoring physiological responses,[76] would be desirable to state agencies in ex-

68. Thomasen, *supra* note 66, at 3.

69. *Id.*

70. *See id.*

71. Jordan Pearson, *The CIA Used Artificial Intelligence To Interrogate Its Own Agents in the 80s*, MOTHERBOARD (Sept. 22, 2014), https://motherboard.vice.com/en_us/article/the-cia-used-artificial-intelligence-to-interrogate-its-own-agents-in-the-80s.

72. *Id.*

73. *Interrogation of an Alleged CIA Agent*, 27 STUD. INTELLIGENCE 49, 53 (1983), https://www.cia.gov/library/readingroom/docs/DOC_0000619182.pdf.

74. *Id.*

75. *Id.* at 54.

76. A human-computer interaction (HCI) system would be equipped with

tracting and evaluating information for truthfulness in interrogations.[77] Robots would not be the first technology incorporated into interrogations to evaluate deception; polygraphs and brain scans have both previously been used.[78] Robots, however, would be equipped with deception-detection technologies and be capable of manipulating conversations.[79] Consequently, they could interrogate and interview better than human counterparts.[80] Such a robot would be a solution to the significant blind spots human interrogators possess when it comes to deception-detection.

c. Interrogation Techniques

Some scholars have argued that robots could meet or exceed the capabilities of the human interrogator, partially because humans are inclined to respond to robots in similar ways as they do to humans.[81] Human-computer interaction (HCI) researcher Joseph Weizenbaum, after studying human conversation with his computer program ELIZA, "concluded that people might even be more willing to open up to automated conversational counterparts than to other humans," particularly if the robot is anthropomorphized.[82]

Robots could be equipped with sensor technology to not only build rapport, but to utilize persuasive techniques as well. These would include flattery, shame, intimidation, and strategic use of body language.[83] Further, designers could use physical subtleties to further personalize the interrogation space, such as manipulating the robot's appearance, voice, and size for strategic purposes.[84] Different attributes, such as "facial structure, voice pitch, tempo, volume, [and] accent," could be manipulated and tailored to perceptions about the individual person

"a combination of visual, auditory, near-infrared and other sensors to monitor a suspect's eye movements, voice, and various other qualities throughout an interaction." Thomasen, *supra* note 66, abstract.

77. *Id.* at 1.
78. *Id.*
79. *Id.*
80. *See id.* at 1–2.
81. *Id.*
82. *Id.* at 4.
83. *Id.* at 1, 3–4, 6.
84. *Id.* at 6.

to be interrogated.[85] Their utility would be in their capacity to be more adept at recognizing human emotions than humans are.

d. Development of Robot Interrogators

Researchers at the University of Arizona developed automated interrogation technology to use in pre-screening interviews of individuals seeking to pass through the United States borders.[86] A member of the team has indicated their usefulness in deception technology: "People have a hard time detecting small changes in the frequency of the human voice, that a computer is much better at. People are accurate about 54 percent of the time at detecting deception . . . . We have got our machine as high as 90 percent in the lab."[87] The Automated Virtual Agent for Truth Assessments in Real Time (AVATAR) is in a testing phase with the Canadian Border Services Agency to assist border security agents in truth-telling determinations.[88] A required kiosk "asks questions of travelers and can detect changes in physiology and behavior during the interview. The system can detect changes in the eyes, voice, gestures and posture to determine potential risk. It can even tell when you're curling your toes."[89] One of AVATAR's developers, Aaron Elkins, claims:

> [AVATAR] can be used not just for border security, but also for law enforcement, job interviews[,] and other human resources applications as well . . . . The system is fully ready for implementation to help stem the flow of contraband, thwart fleeing criminals, and detect potential terrorists and many other applications in the effort to secure international borders.[90]

Another robo-interrogator in development is the more innocuous-sounding Brad at the University of Twente in the Netherlands. Brad is being developed to determine if a robotic

85. *Id.*

86. *Border Patrol Kiosk Detects Liars Trying To Enter U.S.*, HOMELAND SECURITY NEWS WIRE (Aug. 21, 2012), http://www.homelandsecuritynewswire.com/dr20120821-border-patrol-kiosk-detects-liars-trying-to-enter-u-s.

87. *Id.*

88. Suzanne Finch, *The Lie-Detecting Security Kiosk of the Future*, PHYS.ORG (Dec. 28, 2016), https://phys.org/news/2016-12-lie-detecting-kiosk-future.html.

89. *Id.*

90. *Id.*

interviewer can measure physiological cues.[91] Researchers Merijn Bruijnes and Sabine Strofer are conducting experiments to see if robots could measure physiological cues to deception, and if so, whether robots like Brad may be used as police interviewers to fill the demand of governments and agencies in deception-detection.[92]

> [While] Brad is a long way from the kind of fully autonomous police robot that might present a threat to a human's life[,] . . . the use of such techniques in live police operations raises difficult questions, particularly about the potential that a more advanced version of Brad might have for harming its targets.[93]

e. Harm

A significant concern would be the ethical framework in which the robot would operate and its capacity to cause distress.[94] While torture and inhuman treatment during interrogation is prohibited, such rules do not cover or consider robot actors. Indeed, the notorious "enhanced interrogation techniques" used to torture and subject detainees to cruel, inhuman and degrading treatment under the Bush Administration's war on terror could be revived through the moral and ethical issues involved in AI interrogation.[95] "When your captor is a machine, there is no humaneness to be found, and, hence, no one to plead with. When even that small avenue of humanity is done away with in the proceedings of state-sponsored barbarism, what is left? Illegal detainments could continue with only slight human involvement."[96] One example of the potential for distress is Microsoft's AI chatbot Tay:

> [Tay] became a racist monster in fewer than 48 hours, prompting the company to remove the software from the [I]nternet. Unless programmed to work with a highly limited number of interview questions, which could prove too restrictive to be useful, there is potential for any autonomous piece of software to cause distress when interacting with humans, particularly if elevated to a position of institutional authority such as the police.[97]

91. Parkin, *supra* note 67.
92. *Id.*
93. *Id.*
94. *Id.*
95. Pearson, *supra* note 72.
96. *Id.*
97. Parkin, *supra* note 67.

The benefits of robot interrogators have already been recognized. As such, it is imperative to consider the likely presence of robot interrogators in other security contexts as well as how the severing of mens rea and actus reus between a human involved in an interrogation and a robot who engages in unlawful interrogation could evade the CAT's definition of torture and cruel, inhuman and degrading treatment. This growing tear in the fabric of interrogation regulations demonstrates the need for a response in law and policy.

B. LIMITATIONS TO PROTECTION FROM TORTURE COMMITTED BY AUTONOMOUS ROBOTS

The proliferation and projected development of artificial intelligence makes discussion of potential legal challenges, particularly those involving fundamental human rights, of the utmost importance to avoid being "one technological revolution" behind. This Section discusses the foremost treaty on torture prohibitions, the UN Convention Against Torture, the United States' understanding to the CAT requiring "specific intent" for an act to constitute torture, and the subsequent rhetorical evasion that occurred in the infamous U.S. "torture memos" scandal as an illustration of the ways language can be manipulated to avoid obligations under international treaties. Then, it discusses theoretical limitations to protection, specifically cultural approaches to the anthropomorphization of robots and how that may affect language interwoven into incoming visions of technology in the context of a growing post-human movement.

1. The CAT

This Section discusses the exploitable language of the CAT as well as the United States' specific intent understanding to the treaty. Then, it examines the rhetorical evasion in the Office of Legal Counsel's "torture memos" to illustrate the capacity for linguistic manipulation to avoid human rights obligations and justify violations.

a. The Language of the CAT

In 1984, the UN General Assembly adopted the CAT to strengthen existing customary international law norms against torture and require state parties to prosecute anyone who

commits torture within their jurisdiction.[98] Article 1 defines torture as:

> [A]ny act by which severe pain or suffering, whether physical or mental, is intentionally inflicted on a person for such purposes as obtaining from him or a third person information or a confession, punishing him for an act he or a third person has committed or is suspected of having committed, or intimidating or coercing him or a third person, or for any reason based on discrimination of any kind, when such pain or suffering is inflicted by or at the instigation of or with the consent or acquiescence of a public official or other person acting in an official capacity. It does not include pain or suffering arising only from, inherent in[,] or incidental to[,] lawful sanctions.[99]

Article 2 prohibits torture and requires state parties to prosecute or take other effective measures to prevent it in any territory under their jurisdiction.[100] Significantly, Article 2 requires that there are no exceptions to the prohibition, including war, threat of war, public emergency, or command responsibility.[101] Significantly, Article 16 requires parties to prevent, "other acts of cruel, inhuman or degrading treatment or punishment which do not amount to torture as defined in [A]rticle 1."[102]

b. The Nature of Intent

Intent, or *dolus directus*, in international law entails "awareness that by engaging in a certain action or by omitting to act I shall bring about a certain result . . . coupled with . . . the will to cause such result."[103] The majority of international crimes require intent. Intent includes the "natural consequences of [one's] actions" as articulated in the post-World War II *Enigster* case where a Nazi *Schieber* was accused of crimes against humanity for injuries inflicted on Jewish internees in a Nazi concentration camp.[104]

Premeditation is not ordinarily a required element for international criminal responsibility and usually serves as an ag-

98. JEFFREY L. DUNOFF ET AL., INTERNATIONAL LAW: NORMS, ACTORS, PROCESS 413 (3d ed. 2010).

99. Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, art. 1, Dec. 10, 1984, 1465 U.N.T.S. 85, 113–14.

100. *Id.* at 114.

101. *Id.*

102. *Id.* at 116.

103. ANTONIO CASSESE ET AL., CASSESE'S INTERNATIONAL CRIMINAL LAW 43 (3d ed. 2013).

104. *Id.* at 43–44.

gravating circumstance in sentencing decisions.[105] Some international crimes require a special intent, or a *dolus specialis*, for particular crimes, requiring that the "agent pursue a *specific* goal that goes beyond the result of his conduct."[106] These crimes include genocide, persecution, forced pregnancy, and terrorism.[107]

Intent includes recklessness as "a state of mind where a person foresees that his or her action is likely to produce its prohibited consequences, and nevertheless willingly takes the risk of so acting."[108] Recklessness can satisfy the intent requirement for torture and cruel, inhuman and degrading treatment, as exemplified in the 1933 *L. and others* (*Pig-Cart Parade*) case by the Supreme Court in the British Occupied Zone in Germany.[109] In this case, assault troopers paraded a socialist senator and Jewish inhabitant through a small German town, where they were forced to wear demeaning signs around their neck while being led by a pig cart.[110] The defendants participated in vilifying and humiliating actions against the victims, thus:

> The Court held "that, as far as the involvement of three accused went, 'it was inconceivable' that they, who were old officials of the Nazi party, 'did not at least think it possible and consider that in the case at issue, through their participation, persons were being assaulted by a system of violence and injustice; more is not required for the mental element.'"[111]

The International Criminal Tribunal for the Former Yugoslavia (ICTY) developed significant jurisprudence regarding the nature of intent and command responsibility. In *Prosecutor v. Blaškić*, the ICTY Appeals Chamber required "a person 'orders an act or omission with the awareness of a substantial likelihood that a crime will be committed in the execution of that order,' because 'ordering with such awareness has to be regarded as accepting the crime.'"[112]

105. *Id.* at 44.
106. *Id.*
107. *Id.* at 45.
108. *Id.* at 45–46.
109. *Id.* at 47–48.
110. *Id.*
111. *Id.*
112. *Id.* at 48.

c. Rhetorical Evasion: The United States' "Specific Intent" Reservation to the CAT and the U.S. Torture Memos

The United States became a party to the CAT in 1994 and ratified the treaty with various reservations and understandings.[113] Significantly, the United States stated the definition of "torture" required a necessary but not sufficient "specific intent" for an act to constitute torture.[114] Specifically, the understanding stated, "[W]ith reference to Article 1, the United States understands that, in order to constitute torture, an act must be specifically intended to inflict severe physical or mental pain or suffering . . . ."[115] Following the ratification of the CAT, the United States enacted 18 U.S.C. §§ 2340 and 2340A which criminalized any act outside of the U.S. to commit or attempt to commit torture, thereby codifying its treaty obligations in domestic legislation.[116]

Notwithstanding the vast international consensus and expansive legislation,[117] the United States still committed state-sanctioned torture and the Office of Legal Counsel (OLC) justified it with the "specific intent" understanding to the CAT.[118]

Throughout the Bush Administration's tenure,[119] the White House requested legal advice from the OLC to determine if

113. DUNOFF ET AL., *supra* note 98.

114. *See* 136 CONG. REC. S17,486 (daily ed. Oct. 27, 1990).

115. *Id.* (reading of assistant legis. clerk) (unanimous consent agreement and reservations to the ratification of The Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment).

116. 18 U.S.C. §§ 2340–2340A (2012); DUNOFF ET AL., *supra* note 98, at 424.

117. Internationally, the post-World War II Geneva Conventions comprising international humanitarian law as well as the Universal Declaration of Human Rights laying the foundation for international human rights law all explicitly state non-derogable prohibitions against the use of torture. DUNOFF ET AL., *supra* note 98, at 424. Further, the War Crimes Act of 1996 criminalizes violations of Common Article 3 under the Geneva Conventions and the Military Extraterritorial Jurisdiction Act of 2000 further prohibits torture. *Id.* at 434.

118. *See* Jesselyn Radack, *Tortured Legal Ethics: The Role of the Government Advisor in the War on Terrorism*, 77 U. COLO. L. REV. 1, 24 n.161 (2006) (noting the legislative history of the U.S. adoption of the CAT).

119. After the September 11, 2001, terrorist attacks on the World Trade Center towers and the Pentagon, the Bush Administration determined the United States was in a "new kind of war" and that the quick attainment of useful intelligence was critical to the prevention of further attacks and achievement of military success. DUNOFF ET AL., *supra* note 98, at 406. Under

"enhanced interrogation techniques" (EITs), such as waterboarding, would be lawful.[120] Between "2002 and 2007, OLC lawyers John Yoo, Jay Bybee, and Steven Bradbury wrote a series of legal memoranda that served to sanction and authorize what most would deem acts of torture."[121] Significantly, the memos concluded that "a defendant [must] act with the specific intent to inflict severe pain."[122] In order to satisfy the mens rea requirement of specific intent, "the infliction of such pain must be the defendant's precise objective."[123] As such, "even if the defendant knows that severe pain will result from his actions, if causing such harm is not his objective, he lacks the requisite specific intent even though the defendant did not act in good faith."[124] This effectively shielded the prosecution of interrogators who suggested that the obtainment of information and not the infliction of harm undergirded their actions.

The exposure of the secret memos led to a broad array of criticism from international and constitutional law experts for creating loopholes in domestic and international law. Critics claimed the torture memos were designed to provide legal cover for those concerned about potential prosecution for torture and cruel, inhuman and degrading treatment because, under the specific intent standard, "if the accused knowingly causes pain or suffering but had some other objective for which pain and

significant pressure to obtain such information, the Defense Department interrogated many of the thousands of detainees suspected of being terrorists or possessing information about terrorists. *Id.*; Michelle Querijero, Note, *Without Lawyers: An Ethical View of the Torture Memos*, 23 GEO. J. LEGAL ETHICS 241, 242 (2010). At this time, the official interrogation policy in the United States Army Field Manual permitted the use of "psychological ploys" but expressly prohibited "force, mental torture, threats, insults or exposure to unpleasant and inhumane treatment of any kind." DUNOFF ET AL., *supra* note 98, at 406.

120. DUNOFF ET AL., *supra* note 98, at 407.

121. Annamaria Racota, *The Torture Memos: Lawyers, Ethics, and the Rule of Law*, HUM. RTS. BRIEF (Dec. 3, 2009), http://hrbrief.org/2009/12/torture-memos-lawyers-ethics-and-the-rule-of-law. "[T]he torture memos adopted a narrower definition of torture . . . . [A]buse constituted torture only if it resulted in organ failure or death." *Id.*

122. Oona A. Hathaway et al., *Tortured Reasoning: The Intent To Torture Under International and Domestic Law*, 52 VA. J. INT'L L. 791, 793 (2012) (alteration in original) (quoting Memorandum from Jay S. Bybee, Assistant Attorney Gen., to Alberto R. Gonzales, Counsel to the President, on Standards of Conduct for Interrogation under 18 U.S.C. §§ 2340–2340A, at 3 (Aug. 1, 2002) [hereinafter Bybee Memo]).

123. *Id.*

124. *Id.* (quoting Bybee Memo, *supra* note 122, at 4).

suffering was merely incidental, such as extracting information, he lacks the requisite 'specific intent.'"[125] This argument is critical as it lays the foundation for the use of rhetoric to evade liability through the medium of robots and highlights the difficulties in defining intent when the action is splintered between an autonomous agent and its designers, engineers, programmers, and commanders.

While some measures were taken after exposure of these memos,[126] none have served to collapse the gaps capable of being exploited through the use of autonomous weapons in interrogative spaces. Particularly because of the way the OLC utilized rhetoric to claim EITs do not constitute torture if pain and suffering is not the precise objective, uncertainty regarding the types of characteristics we are willing to impute to robots could provide a similarly exploitable opening in intent-analysis. For example, if a robot were to perform an EIT, it would likely be considered to lack the required mens rea due to hesitancy in believing a robot could possess intent. Alternatively, in a more post-human cultural context, a robot may be deemed capable of possessing the required mens rea. Such a determination could make it more difficult to hold human actors involved in the violation criminally liable.

125. *Id.*; *see also* DUNOFF ET AL., *supra* note 98, at 427 (discussing the "necessity" defense to torture).

126. The Senate Intelligence Committee eventually released a report after a four-year investigation concluding the CIA's "use of torture was brutal and ineffective . . . and that the CIA repeatedly lied about its usefulness." Spencer Ackerman et al., *Senate Report on CIA Torture Claims Spy Agency Lied About 'Ineffective' Program*, THE GUARDIAN (Dec. 9, 2014), https://www.theguardian.com/us-news/2014/dec/09/cia-torture-report-released. Further, it was found that the use of torture in the CIA's black sites and prisons was more extreme than previously thought. *Id.* The declassified section of the report reveals many of the EITs used, some of which international law experts say constitute rape under contemporary international law standards. Fionnuala Ní Aoláin, *Sexual Torture, Rape, and Gender-Based Violence in the Senate Torture Report*, JUST SECURITY (Feb. 26, 2015), https://www.justsecurity.org/19731/sexual-torture-rape-gender-based-violence-senate-committee-report. On November 25, 2015, President Obama signed into law an updated National Defense Authorization Act with a provision reinforcing the United States' ban on torture. *President Obama Signs National Defense Authorization Act Solidifying Ban on Torture*, HUM. RTS. FIRST (Nov. 25, 2015), http://www.humanrightsfirst.org/press-release/president-obama-signs-national-defense-authorization-act-solidifying-ban-torture.

2. Theoretical Limitations: Cultural Approaches to Anthropomorphization and Post-Humanism

This Section discusses the significance of cultural approaches to anthropomorphization, focusing specifically on Western and Japanese approaches. Differing tendencies towards anthropomorphization can affect a society's acceptance of a technology and subsequently inform beliefs on liability, responsibility, and acceptance of the emerging technology. Then, it discusses the growing movement of post-humanism and explore how the changing philosophical and rhetorical backdrop to the emergence of autonomous robots could demagnetize linguistic and syntactical aversions of concepts such as "intent," with robots.

a. Cultural Approaches to Anthropomorphization

Technological developments are closely related to and mutually influenced by cultural developments in society, with each level of cultural development enabling new technological innovations.[127] Further, the influence of state, culture, and society can have profound effects on the acceptance of new technologies.[128]

In the United States, fears concerning robots have dated back to the eighteenth century,[129] when, coinciding with the Jean-Jacques Rousseau,[130] the advent of Romanticism, and the

127. *See* N. KATHERINE HAYLES, HOW WE THINK: DIGITAL MEDIA AND CONTEMPORARY TECHNOGENESIS 1 (2012) ("[W]e think through, with, and alongside media."); Volker Boehme-Neßler, *Caught Between Technophilia and Technophobia: Culture, Technology and the Law*, *in* PICTORIAL LAW: MODERN LAW AND THE POWER OF PICTURES 1 (2011) ("New technologies change the world and make people react.").

128. For example, in 1543, Japan acquired firearms, a revolution in military technology for the time. However, in 1607, due to social and political pressure, the Japanese military returned to using traditional weapons, such as swords and lances, and firearms did not return to Japanese warfare until the mid-nineteenth century. This illustration demonstrates the power of culture and law to block progress and even "turn back the clock." Frédéric Kaplan, *Who Is Afraid of the Humanoid? Investigating Cultural Differences in the Acceptance of Robots*, 1 INT'L J. HUMANOID ROBOTICS 465, 472 (2004).

129. *Id.* Kaplan bases this argument on two archetypes: that of an artificially created companion in Greek mythology and the Jewish tradition of alchemy as a way to understand God. *Id.*

130. "To live in civilized societies drives Man far from nature, where he used to live happily . . . . For Rousseau, . . . evolution was the original sin of our species . . . . [S]elf-esteem had replaced the innocent love of our origins and

development of the first automata, the shared idea grew that technical innovations and scientific progress "take man away from his real nature."[131] The nineteenth century saw authors and playwrights expanding on the fear that robots would eventually come to dominate the human race and threaten it with extinction.[132] These fears have also manifested in the belief of the hypothetical event known as The Singularity, in which super intelligent robots ostensibly become capable of recursive self-improvement,[133] leading to a potential runaway effect and intelligence explosion exceeding human intellectual capacity and control.[134] Indeed, in popular American culture, there are several autonomous robots that follow the Western archetype of destroying or seeking to destroy their creator,[135] such as HAL 9000, the antagonist from Arthur Clarke's *2001: A Space Odyssey*,[136] and *Ex Machina*'s Ava.[137] This deeply embedded fear and

the artificial had taken the lead on the natural." *Id.* at 473.

131. *Id.* at 473. "Goethe revived an old Greek tale appropriate for this romantic view of the world: The Sorcerer's Apprentice." *Id.* In the nineteenth century, "Frankenstein syndrome" became a common theme in literature. *Id.* at 474–75. Apart from "the interesting exception of Carlo Collodi's *Pinocchio* (1883), the idea that to create a human-like machine is a transgressive act became common sense." *Id.* at 475.

132. Further, Judeo-Christian monotheism adheres to the doctrine that only God can give life and Exodus decrees idolatry as a sin; in the conventions of science fiction, this usually comes in the form of betrayal of the robots, from *R.U.R.* to *Battlestar Galactica*. *See id.* at 475–76.

133. *See generally* Alva Noë, *The Ethics of the 'Singularity*,*'* NAT'L PUB. RADIO (Jan. 23, 2015), http://www.npr.org/sections/13.7/2015/01/23/379322864/the-ethics-of-the-singularity (providing a variety of perspectives on the concept).

134. Paul G. Allen & Mark Greaves, *The Singularity Isn't Near*, MIT TECH. REV. (Oct. 12, 2011), https://www.technologyreview.com/views/425733/paul-allen-the-singularity-isnt-near (negating the argument it will happen anytime soon).

135. In Japan, popular culture more consistently portrays robots in a positive light. *Better than People: Why the Japanese Want Their Robots To Act More Like Humans*, ECONOMIST (Dec. 20, 2005), http://www.economist.com/node/5323427. *See generally* Kaplan, *supra* note 128, at 466–67 (discussing the cartoon *Testuwan Atom* and exemplifying systematic reintegration of robots into human society as a positive process).

136. *See generally* ARTHUR C. CLARKE, 2001: A SPACE ODYSSEY (1968). HAL is incapable of resolving conflicting programs that include conveying and relaying accurate information, keeping the true objective of the mission a secret, and dealing with the threat of disconnection. *Id.* He kills members of the crew to remove these conflicts. *Id.*

137. EX MACHINA (DNA Films 2014). Ava, a confined humanoid robot with the appearance of a beautiful woman, uses information harvested from the

militarized use of robots could serve to further attenuate the responsibility of a human for the actions of a robot.

These cultural and religious roots of Western notions of the appropriate restrictions on human-computer interaction and the ethics of autonomous robotics can play a role in shaping the way we view autonomous robots' projected capacity for harm as well as their potential legitimate and lawful uses. In comparison with the cultural approaches[138] to robots and the fact that "[America] favors their use [of robots] in war while [Japan] imagines them as benevolent companions suitable for assisting a rapidly aging and increasingly dependent population,"[139] it is possible that some of these assessments are deeply rooted in cultural approaches towards the costs and benefits of the development of systems we can only, for now, imagine, and that these are susceptible to change with the incoming technological revolution.

b. The Radical Repositioning of Post-Humanism on Human-Only Attributes

More concretely, there has been a growing movement of post-humanism that could influence and be influenced by de-

personal information of social media and cell phones to trick the protagonist into helping her escape. *Id.* When she does, she kills her creator and leaves the protagonist to die. *Id.*

138. There is a distinct difference in overall Japanese and American attitudes towards robots. It has been argued that difference is rooted in religion. Christopher Mims, *Why Japanese Love Robots (and Americans Fear Them)*, MIT TECH. REV. (Oct. 12, 2010), https://www.technologyreview.com/s/421187/why-japanese-love-robots-and-americans-fear-them. Japanese social scientist Naho Kitano claims, "Unlike the image of robots of Capek or Asimov, the typical Japan[ese] imagination of robot contains an affinitive rapport between robots and humans. This positive acceptance of robots in the present Japan is founded on Japanese Animism, the idea of *Rinri*, and its rapid modernization." Naho Kitano, Animism, *Rinri*, Modernization; The Base of Japanese Robotics 4 (Jan. 5, 2007) (unpublished manuscript), http://www.roboethics.org/icra2007/contributions/KITANO%20Animism%20Rinri%20Modernization%20the%20Base%20of%20Japanese%20Robo.pdf. Dr. Mitsukuni Yoshida describes the Shinto concept of "anima" as existing within natural and artificial surroundings. *Id.* at 1–2. In the artificial, "anima come alive from the first time as tools or implements that function along with man. And since they are companions of man in life and work, they are often given names." *Id.* at 2. Kitano argues that these religious beliefs combined with the abrupt modernization of Japan led to significant changes in technological development and social systems in a time period too short to allow dynamic changes in ways of thinking. *Id.* at 4.

139. Mims, *supra* note 138.

velopments in the growing autonomy of technology. Advances in technology that equal or surpass those of humans' cognitive or intellectual capabilities may require redefining how we understand what it means to be human. Futurist Ray Kurzweil illustrates this concept in his depiction of a man "frantically scribbling down things that only a human can do" and pinning them to a wall. Slowly, as time passes and technology progresses, the papers start to fall, one by one, to the ground.[140] Since the illustration was first made in 1999, "driving cars," translating "continuous speech," and "cleaning houses" have all fallen to the ground.[141]

The continual redefinition of humanity's place in the world and its exclusive capabilities would hit a cataclysmic crisis upon the development of autonomous robots. Some have argued that thinking beyond humanism will be necessary to redefine humanity's place in the world.[142] Such a radical repositioning could drastically alter our conceptions of intention and responsibility. The difficulties in predicting this only further amplify the need to ensure that an adequate framework for including complicity in torture through the medium of semi-autonomous and autonomous robots is carefully considered and appropriately codified by drafters or experts with requisite levels of technological literacy.

II. OMENS OF POST-HUMANISM THREATENING CORROSION OF TORTURE PROHIBITIONS

This Part examines how a changing rhetorical backdrop may serve to significantly alter understandings of pre-autonomous RMA terms and how the uncertainties stemming from there demonstrate the necessity for a clear framework of liability to encapsulate torture violations. While limited autonomous systems and border avatars demonstrate inchoate autonomy, the explosion of technological advances and resources invested in their development make robot interrogators a very

140. Luke Muehlhauser, Exec. Dir., Singularity Inst., The Singularity, Promise and Peril (2012), https://docs.google.com/document/u/1/pub?id=1BmMESfqoWChb5ZOP85N8Gr72p-Mpv0n8XdbFW3ojWkA.

141. *Id.*

142. CARY WOLFE, WHAT IS POST-HUMANISM (2009) (arguing that humanism, and not humanity, is left behind upon acceptance of posthumanist thought).

real possibility. In this realm of possibility lays a surreptitious threat to our basic fundamental human rights.

Section A explores the exploitable avenues for evasion of culpability in depth, such as "intent" and the splintering of mens rea and actus reus between a human and a robot actor. Section B further focuses on intent, examining how the combination of anthropomorphization, empathy, and post-humanism may collaborate to significantly alter the way we understand "intent" as attributable to non-animal actors.

A. CAN ROBOTS POSSESS "INTENT"?: HOW AUTONOMOUS ROBOTS COULD AFFECT TORTURE PROHIBITIONS

There are differing arguments on robots' capacity for intentionality and "strong" AI. Current opaque understandings are all the more likely to be made even more nebulous upon actual integration of autonomous weapons.

American philosopher John Searle has posited a thought-experiment, "The Chinese Room,"[143] to argue against the idea that a computer could have a "mind"[144] and, consequently, the capacity for intention.

> Imagine a native English speaker, let's say a man, who knows no Chinese locked in a room full of boxes of Chinese symbols (a data base) together with a book of instructions for manipulating the symbols (the program). Imagine that people outside the room send in other Chinese symbols which, unknown to the person in the room, are questions in Chinese (the input). And imagine that by following the

143. Another example of a thought experiment to explain the nature of intent is Gregory S. Kavka's "Toxin Puzzle." *See* Gregory S. Kavka, *The Toxin Puzzle*, 43 ANALYSIS 33, 33–36 (1983). In this thought experiment, you are offered a deal: if you drink a vial of toxin that will make you ill for a day, a billionaire will pay you one million dollars. However, the catch is that you need not drink the vial; you merely need to *intend* to drink it. The money will be deposited into your bank account prior to the time you need to drink the vial. Indeed, "[y]ou are perfectly free to change your mind after receiving the money and not drink the toxin." *Id.* at 34. You desperately try to convince yourself that drinking the vial is a necessary condition in order to meaningfully intend to do so, but you cannot because you know it is not. Kavka uses this illustration to argue that "intentions are better viewed as dispositions to act which are based on *reasons to act* . . . . [Y]ou cannot intend to act as you have no reason to act, at least when you have substantial reasons not to act." *Id.* at 35. This would necessitate the conclusion "that intentions are only partly volitional. One cannot intend whatever one wants to intend . . . . [O]ur intentions are constrained by our reasons for action." *Id.* at 36.

144. *See* John R. Searle, *Minds, Brains, and Programs*, 3 BEHAV. & BRAIN SCI. 417, 417 (1980).

> instructions in the program the man in the room is able to pass out Chinese symbols that are correct answers to the questions (the output). The program enables the person in the room to pass the Turing test for understanding Chinese but he does not understand a word of Chinese.[145]

Since the man could not understand Chinese but was capable of implementing the appropriate program and passing the Turing test,[146] then digital computers also do not "understand" because "no computer, qua computer, has anything the man does not have."[147] In the case of a robo-interrogator, even if it was alone in an interrogation room and implementing a program for torture through the use of its equipped torture armaments, the conclusion does not follow that the robot "understands" such programs *are* torture. Further, even when acting alone in the torture and upon conclusion of the successful extraction of a confession, the robot must be capable of recognizing more than just syntax, as "[t]he purely formal, syntactically defined symbol manipulations don't by themselves guarantee the presence of any thought content going along with them."[148]

There are strong arguments to be made against a robot's capability to possess sufficient intentionality to inculpate them under international human rights law standards.[149] Yet, that would leave an undesirable possibility of human rights violators escaping liability through the use of robots. Further, while robots "have bodies to kick . . . kicking them would [not] achieve the traditional goals of punishment."[150] Yet, if we rule

145. John R. Searle, *Chinese Room Argument*, *in* THE MIT ENCYCLOPEDIA OF THE COGNITIVE SCIENCES 115 (Robert A. Wilson & Frank C. Keil eds., 1999).

146. The Turing test, developed by mathematician Alan Turing, suggested using an "imitation game" to determine if a computer could "think" where "a remote human interrogator . . . must distinguish between a computer and a human subject based on their replies to various questions posed by the interrogator." *Turing Test: Artificial Intelligence*, ENCYCLOPEDIA BRITANNICA, http://www.britannica.com/technology/Turing-test (last visited May 12, 2017). Misidentification as human by the human interrogator indicates a computer's ability to "think." *Id.*

147. Searle, *supra* note 145.

148. *Id.*

149. *See, e.g.*, G.A. Res. 217 (III) A, *supra* note 17; Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, art. 1, Dec. 10, 1984, 1465 U.N.T.S. 85; Rome Statute of the International Criminal Court, art. 28 (entered into force July 1, 2002) [hereinafter Rome Statute].

150. Peter M. Asaro, *A Body To Kick, but Still No Soul To Damn: Legal Perspectives on Robotics*, *in* ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLI-

out this possibility, there still remains no clear answer to the accountability and liability issues for the actions of a robot that are in violation of international human rights law.

However, others have claimed "computers equipped with artificial intelligence (AI) programs" possess intentional systems, where "its behavior is predictable and explainable if we attribute to it beliefs and desires—'cognitive states' and 'motivational states'—and the rationality required to figure out what it ought to do in the light of those beliefs and desires."[151] American naturalist philosopher Daniel C. Dennett illustrates this with the example of Deep Blue, the computer that beat a world chess champion.

> It was Deep Blue's sensitivity to those purposes and a cognitive capacity to recognize and exploit a subtle flaw in Kasparov's game that explain Deep Blue's success . . . [t]he . . . designers of Deep Blue[] didn't beat Kasparov; Deep Blue did. Neither [of his inventors] discovered the winning sequence of moves; Deep Blue did.[152]

As the above examples demonstrate, the nature of intentionality and its attribution to robots is unresolved, intricate, and disputable. Therefore, in the context of an interrogative space, it would likely be as unresolved, intricate, and disputable whether an act of torture committed by a robot met the requisite element of "intent" as stated in the CAT. To wait for the answer to be determined after a violation through a slowly built jurisprudence would be an egregious error if the possibility of preemption existed. Further, a body of case law could move in an unfavorable direction and require significant time and resources to correct initial mistakes. The disputable impact of this concept on intent defeats the state parties' original intent in writing the CAT—to prohibit all torture and cruel, inhuman and degrading treatment, not merely torture by those things philosophically deemed capable of intent.

CATIONS OF ROBOTICS 182 (Lin et al. eds., 2011) (arguing that punishing robots would have little deterrent effect and be comparable to punishing an animal).

151. Dennett, *supra* note 34. While agreeing "skeletal versions of human beliefs and desires [are not] sufficient to meet the mens rea requirement of legal culpability," he states, "it is hard to see what is missing." *Id.*

152. For moral responsibility, rather, he argues it is a "higher-order" intentionality that is required. *Id.* at 354.

B. POST-HUMAN RHETORIC AND THE FLUIDITY OF HUMAN ATTRIBUTES

With the issue of intent unresolved at even the nascent phase in the development of true autonomy, omens of a post-human future threaten to muddy the waters even further. This Section argues that the combination of anthropomorphization of robots, proliferation of robots in domestic spaces, and observed human empathy for robots, all in an increasingly post-humanist cultural mindset, will likely push common understanding of intent towards a belief that robots can and do possess intention. Such an understanding would have undesirable effects on responsibility for human rights violations committed by robots, where those responsible are likely to be viewed as "rogue" or "malfunctioning." This makes it more likely that a complicit person could evade responsibility and redress for victims be prevented.

Computer scientist Noel E. Sharkey has argued that anthropomorphism[153]—the attribution of human traits to non-humans—and "mythical artificial intelligence" have distorted robotics and the public's perception of their capabilities.[154] Such anthropomorphism "plays upon our natural tendency to attribute human or animal properties and mental states . . . to inanimate objects that move in animal-like ways."[155] Sharkey argues that the acceptability of speaking of robots in the anthropomorphic narrative "opens the gates to other meanings associated with the natural language use of the term that may have little or no intrinsic validity to what the computer program actually does."[156]

153. For example, depicting robots as "sentient machines" capable of thoughts, feelings, and desires. Noel E. Sharkey, *The Evitability of Autonomous Robot Warfare*, 94 INT'L REV. RED CROSS 787, 791 (2012).

154. *Id.* at 787 (arguing widespread misunderstandings about limitations of artificial intelligence spread by science fiction and the media and the anthropomorphic ways robots are discussed obfuscates international humanitarian law issues).

155. *Id.* at 791.

156. *Id.* at 793. He argues that "using 'humane' to describe a robot" is like a "Trojan term" that unbeknownst brings with it many other significations. *Id.* Indeed, we are all susceptible to this narrative. Even Gordon Johnson, the Pentagon's former head of its Joint Forces Command stated that robots do not have an appetite, have fear, nor do they forget their orders. Sharkey argues that the fact he felt the need to say those things "[w]ithout being directly anthropomorphic, [he was] leaking it." *Id.* at 792.

Many arguments for the incorporation of robots into the battlefield speak of capacity for being more humane than human soldiers. Yet, some would argue humanity is an "exclusive property of being human."[157] Indeed, the very word "inhuman" is imbedded within the title of the CAT, and considerations as to the "ethical programming" for non-human actors to avoid treating humans inhumanely links human actors and non-human actors as interworking parts of cognition.[158]

These issues may widen rather than collapse the gap in liability. While Searle's thought experiment demonstrates a lack of intentionality due to the absence of a "mind,"[159] the possibility of cultural shifts towards a reorientation of "intent" for the purposes of legislative definitions or common understandings may be underway. N. Katherine Hayles claims we are experiencing the "unfolding story of how a historically specific construction called *the human is giving way to a different construction called the posthuman*"[160] and that "[w]hen information loses its body, equating humans and computers is especially easy, for the materiality in which the thinking mind is instantiated appears incidental to its essential nature."[161]

157. *Id.* at 793. "The statement that robots can be more humane than humans leads to the very worrying implication that robots will humanize the battlefield when in fact they can only dehumanize it further." *Id.*

158. HAYLES, *supra* note 16, at 1.

159. *See* Searle, *supra* note 144, at 424.

160. Ivan Callus & Stefan Herbrechter, *Extroduction: The Irresistibility of the Posthuman: Questioning "New Cultural Theory*," *in* DISCIPLINE AND PRACTICE: THE (IR)RESISITIBILITY OF THEORY 230 (Stefan Herbrechter & Ivan Callus eds., 2004) (quoting N. KATHERINE HAYLES, HOW WE BECAME POSTHUMAN: VIRTUAL BODIES IN CYBERNETICS, LITERATURE, AND INFORMATICS 2 (1999)). It is no coincidence that the changing rhetorical backdrop of society into the post-human world has been reflected in media. The deeply post-human *Westworld* is an early ratings success for HBO and an example of the increased interest in the fledgling philosophy, exploring consciousness of AI and the erasure of the "essential" human. Adam Chitwood, *'Westworld' Ratings Score HBO's Biggest Premiere Since 'True Detective*,*'* COLLIDER (Oct. 3, 2016), http://collider.com/westworld-ratings.

161. HAYLES, *supra* note 16, at 2. N. Katerine Hayles describes the significance of the body in post-humanism as follows:

> [T]he posthuman view thinks of the body as the original prosthesis we all learn to manipulate, so that extending or replacing the body with other prostheses becomes a continuation of a process that began before we were born . . . . The posthuman view configures human being so that it can be seamlessly articulated with intelligent machines. In the posthuman, there are no essential differences or absolute demarcations between bodily existence and computer simulation, cy-

This post-human future implicates the co-evolutionary spiral of technology and humanity as an impetus to reexamination of torture prohibitions' coverage based on humanist conceptions.

As the rhetorical backdrop of various cultures purportedly moves in a post-human direction, human characteristics may continue to become attributed to non-human artifacts with more fluid definitions occurring between the two concepts.[162] Such a cultural shift in post-autonomous RMA terms could dilute a strong case against intentionality and other human characteristics that may seem obvious for now. Such a society has already been realized in science fiction, such as in films *Blade Runner*,[163] *A.I. Artificial Intelligence*,[164] and *Wall-E*.[165] Significantly, recent neurophysiological evidence has emerged demonstrating a human's capacity to feel empathy towards robots in perceived pain.[166] Soldiers in combat have expressed feelings of anger or loss, and some have even held funerals, after the loss of robots that fight with them.[167] Indeed, in the film

bernetic mechanism and biological organism, robot teleology and human goals.

Id. at 3.

162. For example, see Ray Kurzweil's speech on the man in the room with the falling papers. *See supra* Part I.B.2.b.

163. BLADE RUNNER (Warner Bros. 1982); *see also* Nick Lacey, *Postmodern Romance: The Impossibility of (De)Centring the Self*, *in* THE BLADE RUNNER EXPERIENCE: THE LEGACY OF A SCIENCE FICTION CLASSIC (1986) (arguing that *Blade Runner* offers a post-human view of the "replicants," or robots, in its story as manifested by a replicant's sacrificial behavior and realization of self-worth transforming him into a human). Yet, deeply embedded cultural narratives in the West about the destructiveness that comes from "creating life" are also present in tales of caution regarding these stories. For example, in *Ex Machina*, an android called Ava is programmed to attempt to escape her confinement in a room. EX MACHINA (DNA Films 2015). She is allowed contact with the protagonist, an outsider, who is unaware of her programming and her intent, and convinces him that she does possess these human attributes and characteristics. *Id.* Upon release, she kills her creator and leaves the protagonist to die, demonstrating the complete feigning of emotion in accordance with achieving her programmed objective. *Id.*

164. A.I. ARTIFICIAL INTELLIGENCE (Warner Bros. 2001).

165. WALL-E (Disney 2008).

166. *Humans Can Empathize with Robots*, SCI. DAILY (Nov. 3, 2015), http://www.sciencedaily.com/releases/2015/11/151103064554.htm.

167. The robots were also given names, usually after wives or girlfriends, and the men and women actually interacted with the machines more like pets than metal. Meghan Neal, *Are Soldiers Getting Too Emotionally Attached to War Robots?*, MOTHERBOARD (Sept. 18, 2013), https://motherboard.vice.com/

Star Wars Episode VI: Return of the Jedi, the viewer is meant to feel empathy for a robot being tortured in a gangster's lair, who audibly screams as his "feet" are being burned.[168]

Do these elements add up to an overarching understanding of the potential for sentience, and subsequently, intentionality in robots? Possibly. The proliferation of robots, particularly in domestic spaces, is more likely to contribute to this belief. However, such a conclusion can reflect the erosion of the CAT's language and undermine its purpose to hold humans responsible for violations of non-derogable human rights norms. Such a cultural response to an increasingly shared world with robots made in likeness and image to humans only further necessitates a reexamination of the language in the CAT to ensure liability for those responsible for torture.

C. HOW PRESENT THEORIES OF LIABILITY FAIL TO REACH INTO A POST-HUMAN FUTURE

Autonomous weapons are a much-debated topic and a significant number of arguments have been made advocating for effective regulation and liability frameworks. However, many of these frameworks may be rendered precarious upon the converging boundaries of autonomous weapons and post-human understandings of liability. The changing landscape autonomous robots will likely evoke upon their inhabitation threatens product liability and command responsibility frameworks with stagnation. Indeed, "[t]he mere fact that a human might not be in control of a particular engagement does not mean that no human is responsible for the actions of the autonomous weapon system."[169]

This Section discusses the theories of accountability concerning human responsibility for the actions of robots, focusing on product liability and command responsibility. It then evaluates why they are not sufficient to prevent human rights viola-

en_us/article/are-soldiers-getting-too-emotionally-attached-to-war-robots. After the destruction of one soldier's robot, he "brought in [the] robot for repairs with tears in his eyes and asked the repair shop if it could put 'Scooby-Doo' back together. Despite being assured that he would get a new robot, the soldier remained inconsolable. He only wanted Scooby-Doo." *Id.*

168. STAR WARS: EPISODE VI – RETURN OF THE JEDI (Lucasfilm 1983).

169. SOLIS, *supra* note 14, at 544 (quoting Michael N. Schmitt & Jeffrey S. Thurnher, *"Out of the Loop": Autonomous Weapons Systems and the Law of Armed Conflict*, 4 HARV. NAT'L SECURITY J. 231, 277 (2013)).

tions without an additional specific prohibition on torture through the medium of autonomous robots.

1. Product Liability: Engineer, Designer, and Programmer Responsibility

Because criminal liability is predicated on a guilty state of mind, the lack of a "mind" in the agent makes imputation of liability to operators or designers difficult to transfer and therefore, assess.[170] Some have argued against the illusion that increasing autonomy correlates with a diminished responsibility of the designer and operators for the artificial agent's acts.[171] Instead, increasing autonomy would add more to designers', operators', and programmers' burdens.[172] As robots encounter challenges, particularly those on the margins of predictability, human operators will need to anticipate what the robot may do—a daunting task, particularly as systems become more complex and autonomous. If we attribute a robot's failure to their operators, "to expect operators to anticipate the actions of intelligent systems becomes more and more unreasonable as the systems and the environments in which they operate become more complex."[173]

Indeed, an autonomous robot in an interrogative space self-sufficiently conducting an interrogation and utilizing torture—already a much disputed space[174]—invokes an array of issues, including the disputed lines on what constitutes lawful interrogation, enhanced interrogation, and torture or cruel, degrading and inhuman treatment. The expectation of designers to predict challenges a robot may encounter in these margins during an interrogation task is daunting. As complexity grows and autonomy evolves, it may become "even less realistic to expect human operators to exercise significant veto control over their operations . . . . [I]t would be unfair to hold an individual operator accountable for failing to hit the 'off' switch."[175]

170. *See* Grut, *supra* note 4, at 16; Searle, *supra* note 144.

171. Wendell Wallach & Colin Allen, *Framing Robot Arms Control*, 15 ETHICS INFO. TECH. 125, 131–32 (2013).

172. *Id.*

173. *Id.* at 132.

174. *See infra* Part III.A.

175. Grut, *supra* note 4, at 15.

Further, to impute responsibility and accountability to engineers and designers invokes a product liability framework. Yet, it is difficult to see "a monetary fine that the firm's insurance company will end up paying"[176] as an equitable punishment for a violation of a jus cogens human rights norm. Rather, punishments should be aimed at preventing the recurrence of the violation and ensuring the right to be free from torture or cruel, inhuman and degrading treatment for all.

2. Command Responsibility: Commanders, Senior Staff, and Policymakers

Another theory has been to apply the doctrine of command responsibility from Article 28 of the Rome Statute of the International Criminal Court to autonomous weapons. Article 28 states:

> (a) A military commander or person effectively acting as a military commander shall be criminally responsible for crimes within the jurisdiction of the Court committed by forces under his or her effective command and control, or effective authority and control as the case may be, as a result of his or her failure to exercise control properly over such forces, where:
>
> (i) That military commander or person either knew or, owing to the circumstances at the time, should have known that the forces were committing or about to commit such crimes; and
>
> (ii) That military commander or person failed to take all necessary and reasonable measures within his or her power to prevent or repress their commission or to submit the matter to the competent authorities for investigation and prosecution.[177]

While it seems clear that the Rome Statute refers to people under the term *forces*, it is not unreasonable to include robots under that term, particularly if the robot is replacing a "human role of the soldier."[178]

The Director for International Law and Cooperation at the International Committee of the Red Cross (ICRC) has stated that "someone has to be responsible for turning the system on."[179] Further, a system needs to have some degree of predictability to ensure that a commander would be in a position where

176. *Id.* at 16 (quoting P.W. SINGER, WIRED FOR WAR: THE ROBOTICS REVOLUTION AND CONFLICT IN THE 21ST CENTURY 412 (2009)).

177. Rome Statute, *supra*, note 149.

178. Grut, *supra* note 4, at 18; *see also id.* at 18 n.64.

179. *Id.* at 19.

he "should have known"[180] that the system was committing or was going to commit a crime. However, this would entail a decision between the commanders who gave the order to use the robot as the last point of contact or the senior staff and policymakers who made the decision to deploy the robot.[181] Notwithstanding this tension, the issue of predictability remains. Particularly due to the complexity of the development of such systems, there is:

> [A] common misconception that robots will do only what we have programmed them to do . . . . [P]rograms with millions of lines of code are written by teams of programmers, none of whom knows the entire program; hence, no individual can predict the effect of a given command with absolute certainty, since portions of large programs may interact in unexpected, untested ways . . . .[182]

In addition, there remains the question of whether autonomous robots would even obey orders or be capable of recognizing a chain of command.[183]

In the context of the IT-O, a robot clearly designed for torture and interrogation, the way that liability should be distributed among the commanders to the officers in charge of interrogation to the designers of the system may seem a bit clearer. However, if the system is not, or is not clearly, designed for torture and has other legitimate purposes in lawful interrogation, like Brad, yet still commits torture or inhumane treatment outside of its intended purpose, the spread of liability becomes more difficult to assess. Further, issues on the undulating waves of increasing complexity in programming, robo-human relationships, and integration into hierarchical structures, call into question these theories' sustainability into a more fully realized autonomous and post-human future.

The combination of the continuing anthropomorphization of robots along with the potential for a more widespread post-human rhetoric that increases fluidity of human characteristics among bodies as well as attributes could further widen the gap of liability to be applied to those complicit in torture and cruel, inhuman and degrading treatment. It opens the possibility for

180. Rome Statute, *supra*, note 149.

181. Sharkey, *supra* note 153, at 790–91.

182. Grut, *supra* note 4, at 20 (quoting Gary E. Marchant et al., *International Governance of Autonomous Military Robots*, 12 COLUM. SCI. & TECH. L. REV. 272, 283 (2011)).

183. *Id.*

a deeper trench separating the robots, their commanders, and their creators, where those robots who go too far in an interrogation are "rogue" or "malfunctioning" rather than a symptom of complicity. As such, it is absolutely pivotal to collapse this gap preemptively. Without doing so, as robots are increasingly incorporated into interrogation,[184] there remains the possibility of heinous human rights violations.

III. DRAFTING AN ADDITIONAL PROTOCOL TO THE CAT TO SPECIFICALLY BAN TORTURE THROUGH AUTONOMOUS ROBOTS

Part II of this Note discussed how product liability and command responsibility frameworks fall short due to under-inclusiveness and over-inclusiveness of liability as well as for failing to catch the precise responsible actors in its netting. An underlying issue is the uncertainty of the equitable nature of responsibility for predicting a robot's behavior in conjunction with increasingly complex programming. As such, as the technology develops and robotics engineers increasingly develop robots capable of near-true or true autonomy, there can be no one-size-fits-all sustainable framework. Rather, we must engage in a continuous analysis of prediction, prevention, and retailoring to accurately reflect the issues of torture through mediums of robotics.

Additional Protocols are convened when the landscape of the space a treaty sought to cover has changed so drastically that new reformulations are required to encapsulate the space's undulations through time. The two Additional Protocols to the 1949 Geneva Conventions were adopted by states to codify and implement international humanitarian law more completely and universally to modern armed conflicts.[185] This update was necessary to cover the unforeseen evolution in armed conflict. While the Geneva Conventions "looked back" on the Second World War and interstate, international armed conflict waged between belligerent states, the Additional Protocols sought to expand civilian protection in the midst of guerrilla intrastate

184. Neil M. Richards & William D. Smart, *How Should the Law Think About Robots*?, *in* ROBOT LAW, *supra* note 6, at 1.

185. *Protocols I and II Additional to the Geneva Conventions*, INT'L COMMISSION RED CROSS, https://www.icrc.org/eng/resources/documents/misc/additional-protocols-1977.htm (last visited May 12, 2017).

wars with non-state actors as parties to the conflict. Following the logic of actualizing the living intent of the drafters of the 1949 Geneva Conventions, so too will the CAT require updates to supplement the protections of the original treaty while better carving out protections for the future based on the evolutionary path of interrogation spaces.

The Additional Protocol to the CAT would need to establish an initial framework of responsibility to appropriately include those actors who would be equitably responsible for the actions of the robot based upon the present and predicted capacity of robots for the near future. "[T]he law should be carefully calibrated to express his or her exact level of culpability"[186] and to do so is highly dependent on the space in which an autonomous robot is used. While there has been significant discussion regarding liability and responsibility for the use of robots in war crimes spaces in violation of international humanitarian law principles, discussions have centered less on robo-interrogation.

This Part examines guidelines for drafting the proposed Additional Protocol for regulating autonomous weapons in interrogative spaces. It argues that utilizing an autonomous robot in an interrogative space could foreseeably result in a human rights violation based on the intricacies of the lines between torture and cruel, inhuman or degrading treatment and lawful interrogation. In this context, guidelines include revisiting and revising terms, conduct-based syntactical composition choices, coverage of threat to torture, and digital literacy of the drafters. It also discusses critiques of the arguments made in this Note in support of an Additional Protocol.

A. CONSIDERATIONS FOR DRAFTING TERMS OF THE ADDITIONAL PROTOCOL

This Section addresses five guidelines for the drafting of an Additional Protocol to the CAT. First, it goes through a required reconvening period for continual reanalysis based on the technological and cultural climate and second, a linguistic differentiation of humans and robots based on actual terms and conduct-based analysis rather than perceived human-only attributes and intent-based analysis. Then, it suggests a foresee-

186. Jens David Ohlin, *The Combatant's Stance: Autonomous Weapons on the Battlefield*, 92 INT'L L. STUD. 1, 30 (2016).

ability framework to consider in an eventual regulation of liability. Further, it recommends considering how robots' presence can be used as a mere threat to torture and the need to cover that potential space.[187] Last, it suggests a requisite level of "electracy."[188] These guidelines are aimed as a supplement to an eventual theory of liability to consider language framed to ensure that it is not frozen in time and prone to rapid decay.

1. Required Reconvening for Continual Reanalysis

Critical to the Protocol would be a renewal period. If the state of technology has not advanced so dramatically that the existing framework still works in an effective manner, the framework must be renewed. If, however, it has been determined that there are dramatic advancements on the spectrum of autonomy, then state parties must reconvene to discuss advancements in frameworks to determine the appropriate scope based upon new considerations of the human-robot relationship, its prevalence and predictability, and the nature of complicity. These would reassess the language used and frameworks of liability to fluidly evolve alongside technology. Such a requirement could evade rendering the CAT and its Additional Protocol obsolete, especially in the context of an ever-evolving technological world.[189] While such a renewal period could be criticized as a waste of resources, unprecedented, and inconvenient, it is critical for an exponentially proliferating technology

187. Gäfgen v. Germany, 2010 Eur. Ct. H.R. 1.

188. Kathryn Barrett, *What Is Electracy?*, (Oct. 9, 2009), https://mpm17fall2009friday.wordpress.com/2009/10/09/what-is-electracy-3 (explaining "electracy," the term coined by Gregory Ulmer, "relates to the idea of being literate in a digitally mediated world").

189. *See Moore's Law*, ENCYCLOPEDIA BRITANNICA, http://www.britannica.com/topic/Moores-law (last visited May 12, 2017). Moore's law predicted that "the number of transistors per silicon chip [or more generally, computing power] doubles every year" (adjusted to every two years), leading to an explosion in complexity and exponential growth. *Id.* The prediction has been "impressively true," and is often used for the proposition that technology advances with exponential growth. Graham Templeton, *What Is Moore's Law?*, EXTREME TECH (July 29, 2015), http://www.extremetech.com/extreme/210872-extremetech-explains-what-is-moores-law; *cf.* Chris Green, *The End of Moore's Law? Why the Theory That Computer Processors Will Double in Power Every Two Years May Be Becoming Obsolete*, INDEPENDENT (July 16, 2015), http://www.independent.co.uk/life-style/gadgets-and-tech/news/the-end-of-moores-law-why-the-theory-that-computer-processors-will-double-in-power-every-two-years-10394659.html.

qualitatively different than those of the past. The renewal period provides room to address and confirm the effectiveness of coverage and also provides for an ability to reconvene and adjust without requiring further treaties and Protocols for eventual changes in linguistic, psychological, and integrative command realities in the way we relate to and interact with autonomous robots.

2. An Unequivocal Ban: From Intent to Conduct-Based Analysis

The language of the ban should utilize specific references to humans rather than to perceived human-only attributes. As Ray Kurzweil's wall demonstrates,[190] defining humans by what "only humans can do" is merely taking a snapshot in time and holding it as a universal truth. Little by little, conceptions of what it means to be human may disintegrate and leave the prohibition incomplete with ambiguous application. Rather, an Additional Protocol should unequivocally state that the use of robots in interrogative spaces by state actors may foreseeably result in human rights violations of which human indirect perpetrators will be held liable.

In this vein, the basis on evaluating whether an act constitutes torture should be a conduct-based rather than an intent-based analysis. The "torture memos" demonstrated the perversion of the language in the CAT contrary to the drafters' intent.[191] In the Additional Protocol, shifting the definition of torture to conduct rather than intent would better reflect the drafters' intent and victims' rights.

3. Utilizing Growing Foreseeability Jurisprudence

Further, similar to the growing jurisprudence of foreseeability of harm in conflict spaces, the Additional Protocol should consider using a robot in an interrogative space in combination with other elements foreseeable to commit a human rights violation. For example, as more is learned about conflict spaces,

190. *See also* Henry Nicholls, *Jane Goodall: How She Redefined Mankind*, BBC (Apr. 1, 2014), http://www.bbc.com/future/story/20140331-the-woman-who-redefined-mankind (discussing how Goodall's discovery that chimps could use tools—an ability previously only attributed to humans—changed our definitions of what makes humans unique).

191. *See supra* Part I.B.

certain harms are considered foreseeable given combinations of differenct factors. For example, indicators such as displacement, forced conscription, and discrimination against women can contribute to a finding that sexual violence is foreseeable in conflict.[192] This growing jurisprudence can be instructive for robots in interrogative spaces. As holistic narratives of interrogation and autonomous robots become more present, we can understand and address what factors increase the risk.

In addition, there lies the potential for proposing to insert new types of entities into the interrogative space to be able to carry out something an individual, as a human interrogator, knows they would be unable to do themselves. Using foreseeability jurisprudence can help to collapse this potential for exploitation.

4. Threat of Torture

Another consideration in the drafting of an Additional Protocol would be to consider how a robot's presence could be used as threat of torture. Although the legality of threat of torture has not been settled in international law,[193] the presence of a robot in an interrogation space may lead a detainee to believe that a human may be the sole source of mercy. One can easily see how that dynamic could imply impending torture and even be used purposefully to do so. As such, an Additional Protocol should unambiguously address this situation, as well.

5. Considering "Electracy" & Digital Literacy

Lawmakers or legal advisers should be literate in these kinds of discourses to intelligently and appropriately codify regulations on emerging technologies given their interwoven existence with human cognition. Potentially applicable here is the theory of "electracy," explained as "to digital media what literacy is to print."[194] Interactions with and between technolo-

192. Barbara Goy et al., *Contextualizing Sexual Violence and Linking It to Senior Officials: Modes of Liability*, *in* PROSECUTING CONFLICT-RELATED SEXUAL VIOLENCE AT THE ICTY 250–54 (Serge Brammertz & Michelle Jarvis eds., 2016).

193. *See* Gäfgen v. Germany, 2010 Eur. Ct. H.R. 1. *See generally* SOLIS, *supra* note 14, at 437–85 (discussing how the laws around threat of torture are unsettled).

194. *See* Barrett, *supra* note 188.

gy, coding and programming should entail a wider understanding of those processes and their digital presentations in order to effectively capture specific behavior.

B. CRITIQUES

There are several potential critiques for the arguments advanced in this Note. The first is that post-humanism is a uniquely and isolated, culturally specific phenomenon and this problem may not look similarly to countries or cultures without access to advanced robotics technology. As a cultural rhetoric produced by the evolution of the very society that creates it, to consider post-humanism in the context of autonomous interrogation when choosing the drafting language of a regulation would be to cater solely to states experiencing the societal change brought on by proliferation of robots. Further, the convening and rulemaking would be undertaken by states with asymmetric access to autonomous robots.

In response to this, it is important to recognize that international agreements are created out of an amalgam of interests. It is not an anomaly for new changes or adjustments to be undertaken because of the need of a minority of country interests.[195] Such an agreement would not only be in the interest of countries with the technology, but those without, as well. Significantly, to only allow those countries with currently projected possession of such technologies excludes countries that may wish to protect their nationals and troops in armed conflict and otherwise from such interrogation techniques. This new dilemma is not actually new and mirrors the development of the 1899 Hague Regulations, wherein Russia, concerned about the rapid industrialization of the West, wanted protection from advancing technologies.[196] Access to a "seat at the table" permits blocks of countries without access to such weapons to negotiate and establish their positions in a meaningful way.

195. For example, many attribute the impetus for creation of the Rome Statute to Trinidad and Tobago's interest in illicit drug trafficking. *See Overview*, *Rome Statute of the International Court*, UNITED NATIONS, http://legal.un.org/icc/general/overview.htm (last visited May 12, 2017) ("In December 1989, in response to a request by Trinidad and Tobago, the General Assembly asked the International Law Commission to resume work on an international criminal court with jurisdiction to include drug trafficking.").

196. *See* SOLIS, *supra* note 14, at 51 n.62 (citing the Hague Conventions and Declarations of 1899 and 1907).

An additional counterargument is that the CAT's language is sufficient to cover potential violations by robots in interrogative spaces. An accurate rebuttal to this critique goes beyond the scope of this Note insofar as it implicates the larger debate regarding the efficacy of the CAT to cover the broad spectrum of interrogation and torture spaces and potential harms that exist.

The arguments in this Note do in fact fit into the broader argument that the CAT in its current form is insufficient to deal with human interrogation techniques. If the CAT cannot sufficiently address, prevent, and bind human interrogators of states, it is unrealistic to rely on this framework to also cover robot interrogators. On an operational level, the language of the CAT was insufficient to prevent lawyers from perverting its text, as illustrated by the OLC "torture memos" and seven years of a comprehensive CIA torture program.[197]

Further, sexual violence and gender-based violence also expose some of the insufficiencies of the CAT to cover all torture spaces and implications of some coercive interrogation techniques. Scholars are considering the ways in which cultural heritage, religious beliefs, and gender can be exploited in interrogation spaces and can amount to torture.[198] These variables are not clearly addressed in the CAT and yet, are standardized or even praised as innovative interrogation methods. In sum, if the CAT cannot address the ways in which states compel their human interrogators, one cannot expect the CAT to dissuade or prevent states from giving innovative instructions to evade the CAT or from using non-humans to undertake similarly degrading techniques.

Last, it could be argued that we are better served by an entire convention on autonomous weapons as a whole rather than an Additional Protocol to uniquely address the question of robots in interrogative spaces. Such a convention could address the various facets and spaces autonomous robots could affect,

197. *See supra* Part I.B.

198. *See, e.g.*, Ramzi Kassem, *Gender Erasure in the Global "War on Terror," in* GENDER, NATIONAL SECURITY, AND COUNTER-TERRORISM: HUMAN RIGHTS PERSPECTIVES 24 (Margaret L. Satterthwaite & Jayne C. Huckerby eds., 2013); Olivier Winants, *The Interplay of Ethnicity, Gender and Sexual Violence During Wartime and in Coercive Interrogation: What Role for Human Dignity?*, 43 JURA FALCONIS 137, 137–74 (2006).

rather than inserting additional provisions or protocols into already established treaties.

Although the desire for an entirely new convention may be well founded, given contemporary conflict and security spaces and rising political trends of isolationism and anti-globalization, it is something unlikely to occur. However, an Additional Protocol would serve as an update to something states have already widely agreed to in order to better reflect new and emerging technologies. This space is specific to interrogation and must be contextualized within the foremost treaty on torture. It would be more efficient to include the prohibition where it is most relevant rather than couching it in the expansive, multi-faceted, and potentially problematically vast issues faced upon introduction of autonomous weapons. Further, the Additional Protocol would provide an opportunity to update the CAT in light of emerging norms in international human rights law, including significant cultural and religious considerations of interrogation spaces to expand the CAT's coverage in a more meaningful and comprehensive way.

CONCLUSION

As we approach an impending technological revolution and proliferation of robots on the spectrum of autonomy, we run the risk of being "one technology behind" in anticipating the changing landscape in the next season of human-technology interaction. Autonomous robots have a unique capacity to splinter a criminal act, where a human manifests the mens rea and the robot commits the actus reus. Such a possibility poses serious challenges to the current international frameworks for imputing liability for complicity in torture, especially if robo-interrogation continues down its projected evolutionary path. Perceptions regarding what it means to be a human and what characteristics, beliefs, intentions, and desires belong exclusively in the realm of the human and which will "fall from the wall" of humanity into the shared realm of the post-human will be challenged. Such potential rhetorical evolutions could stagnate the current use of the CAT's language of "intent." Grasping into the future for significations around a technology that does not yet exist can feel like writing in the sand only to have the tide wash it away. Yet, robots are already in their infancy, and to ignore the likely growth of their lifespan would be to re-

peat fatal mistakes of the past and other nascent, deadly technologies.

This Note posits an Additional Protocol to the CAT to close in on the exploitable gaps upon integration of robots and a posthuman world. This Note has provided guidelines to the drafting of an Additional Protocol so as to sew back together actus reus and mens rea initially splintered through use of an autonomous robot as a medium to commit torture. The Protocol should require a reconvening and renewal period so as to reassess the language used and frameworks of liability in order to fluidly evolve as the technology does. As robots are incorporated into and effect change upon the world, such guidelines are necessary to ensure the dignity and fundamental human rights of all persons to be free from torture and cruel, inhuman and degrading treatment.